Functional MRI: Data Analysis, Reproducibility, Reliability and Pitfalls

Peter Jezzard, PhD

FMRIB Centre, John Radcliffe Hospital, Headington, Oxford OX3 9DU, England

Introduction

Functional MRI has seen an explosion in its growth since its earliest human demonstration in 1992 (Fig. 1). The technique has been used predominantly to date as a method for studying the functional organisation of the healthy brain. As the subject matures, however, fMRI is increasingly being used to study clinical populations in areas of neurological and psychiatric disease. For example, functional MRI offers the possibility of contributing to clinical study by mapping normal brain function in relation to adjacent lesions (e.g., tumour resection and epilepsy surgery), for assessing abnormal oxygen metabolism, for studying disease status and progression, and as a marker for drug action and therapeutic efficacy in drug development. However, there are still considerable uncertainties in the robustness of applying functional MRI in the presence of patho-physiology, and there is still a need to have inter-center standards of fMRI paradigm implementation, data analysis, and scanner reliability before fMRI can be effectively used as a clinical tool.

Principles of Functional MRI

The human brain is composed of approximately 100 billion neurons, each with between 1,000 and 10,000 connections. The total number of interconnections in the brain is therefore on the order of 10^{14}–10^{15}. Communication amongst neurons occurs between the pre-synaptic terminal (on the axonal side) and post-synaptic membranes (on the dendritic side). Surrounding the neurons are glia, which are cells involved in energy metabolism, storage, and maintenance of ionic balance.

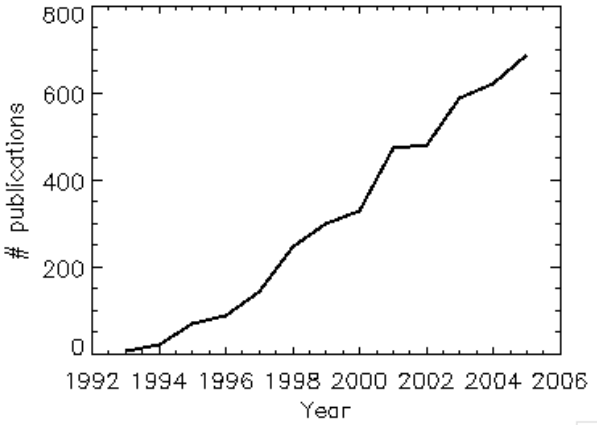

Fig. 1: Annual number of publications listed in PubMed with either "fMRI", "functional MRI" or "functional magnetic resonance imaging" in the title, since 1992.

The metabolic turn-over of glia, activation of neurons, and establishment of the ion potentials in the cells of the brain all require a supply of energy. This energy is provided in the form of adenosine tri-phosphate (ATP) generated in the mitochondria within cells. Current opinion is that the most energy consuming processes in the human brain are the post-synaptic processes involved in neurotransmission. Hence, fMRI appears to report mostly on input of information into a brain area. Under normal conditions ATP is formed via glucose consumption in the presence of oxygen (via aerobic glycolysis), and this glucose and oxygen is supplied by blood perfusing the tissue. Previous experiments involving a number of modalities have shown that local glucose consumption rises sharply when neuronal activation takes place. This is accompanied by an increase in local blood flow, and in local blood volume. From the perspective of blood oxygenation level dependent (BOLD) fMRI, however, it is the balance of oxygen delivery and utilization that results in the observed signal changes. This is because blood oxygen is predominantly transported within red blood cells, bound to the large iron-containing molecule, hemoglobin. It is the changing magnetic properties of hemoglobin

as it gives up its oxygen that provides the most utilized contrast mechanism in functional MRI. Other physiological parameters that can be accessed by MRI include cerebral blood flow (Wong *et al*., 2000), cerebral blood volume (Lu *et al*., 2003) and, in animal models, glucose turn-over via carbon-13 spectroscopy (de Graaf et al. 2004).

The particular origin of the BOLD fMRI effect lies in the different magnetic properties of hemoglobin that has oxygen attached to it (oxyHb) versus hemoglobin that does not have oxygen attached (deoxyHb). A magnetic susceptibility difference of approximately 0.08 ppm is observed between fully deoxygenated whole blood and fully oxygenated whole blood. Since fully oxygenated blood is isomagnetic relative to tissue, vessels containing arterial blood cause little or no distortion to the magnetic field in the surrounding tissue. Conversely, capillary and venous vessels containing blood which is partially or significantly deoxygenated will distort the magnetic field in their vicinity (see Fig. 2).

The NMR signal *change* that is measured is composed of contributions that originate from two compartments. These are: (i) the water molecules in the blood itself (intra-vascular compartment), which in turn have a pure T2 contribution and a T2* contribution, and (ii) a contribution that originates from water molecules in the tissue space that surrounds the vessels (extra-vascular compartment). Depending on the pulse sequence that is used, a different relative signal change from each compartment can result. The BOLD signal that is observed will be a volume-weighted average of the intra-vascular and extra-vascular water signal originating from within and around the capillaries and veins (Boxerman *et al*. 1995).

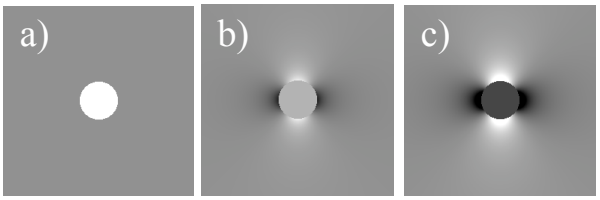

Fig. 2: Field inhomogeneities created around a vessel orientated parallel to the B_0 field (a), at 45º to the B_0 field (b) and at 90º to the B_0 field (c).

Data Analysis Issues

Analysis of functional MRI data has become a sophisticated art, and there are a number of excellent software packages available either commercially or as freeware. The main steps in the analysis pipeline involve the following:

i) Correction for motion mis-registration. Often the subject will move slightly during the course of data collection. This will result in artifactual signal intensity variations in the pixel time course that are attributable to changes in the tissue content and location within the pixel, rather than to changes in deoxyHb content. Usually a rigid body motion correction is used to correct any mis-registration (e.g. Woods *et al*., 1998; Jenkinson *et al*., 2002). Strictly the use of a rigid body algorithm may not be valid in the case of echo planar imaging, where interactions can occur between the motion and the inherent geometric distortion of EPI (Hutton *et al*., 2002). Also, additional care should be taken whenever the head motion correlates with the applied stimulus (Friston *et al*., 1996).

ii) Slice timing correction. Typically fMRI data are collected using a multi-slice EPI sequence that necessarily collects the individual slices within a given volume at different time points. As such, the hemodynamic model that is used to correlate to the empirical data will not match the temporal profile of every slice unless a timing correction is applied. This can be done either by temporal interpolation of the image data, or by temporally shifting the model.

iii) Pre-statistics. It is often advantageous to apply spatial and temporal filtering of the time course data, to better match the characteristics of the data to the known limits of the hemodynamic response. Usually a band-pass temporal filter is applied to remove both slow drifts (often residual motion artefacts or scanner drift problems), and to remove high frequency fluctuations (often aliased cardiac and respiratory-related effects). Note that the characteristics of the low-pass side of the filter must not interfere with the frequency of the applied stimulus paradigm. A modest amount of spatial smoothing is also often applied to improve the signal-to-noise ratio of the data and to enforce Gaussian spatial properties in the data.

iv) Construction of a model. A theoretical model is constructed to represent the expected time course of the presumed electrical activity in the brain. This is convolved with an estimate of the hemodynamic response function, in order to obtain a theoretical pixel time course for an area of the brain that is engaged in the task under study. Potentially, several such linear models can be constructed in the form of a 'design matrix'

v) Statistical analysis of the data. The theoretical model is then fitted to the experimental data by least squares methods. This results in a set of parameter estimates for the amplitude of each linear time course (also known as explanatory variables) in the model, plus the residual noise variance. From this a T-statistic can be calculated at each pixel location by dividing the size of the parameter estimate by the residual noise estimate from the model. A T-map to Z-map transformation may then be applied.

vi) Statistical thresholding. The Z-maps are then thresholded according to a statistical confidence level that is dictated by the experimenter. This should ideally be high enough to avoid false positives, but low enough to avoid false negatives.

vii) Structural registration. It should be noted that many function MRI acquisition sequences suffer from geometric distortion when compared to high resolution standard structural sequences. This is particularly true of echo planar imaging, although with field map information it is possible to correct much of this distortion (Jezzard and Balaban, 1995). Clearly, such problems are of particular importance in pre-surgical mapping.

There are several issues that pertain when analysing data from clinical populations. First, traditional fMRI studies often ultimately rely on a group analysis in order to answer a particular hypothesis. In clinical studies this may or may not be the case, depending upon whether the study is part of a clinical research question (perhaps involving the comparison of a group of patients with a group of control subjects), or is a specific application in a single clinical subject (perhaps for pre-surgical mapping). In the latter case it is essential that rigorous data analysis be performed in order to avoid potential confounds. For example, the statistical thresholds that may be suitable for group analyses (where false negatives may not be overly problematic) may be

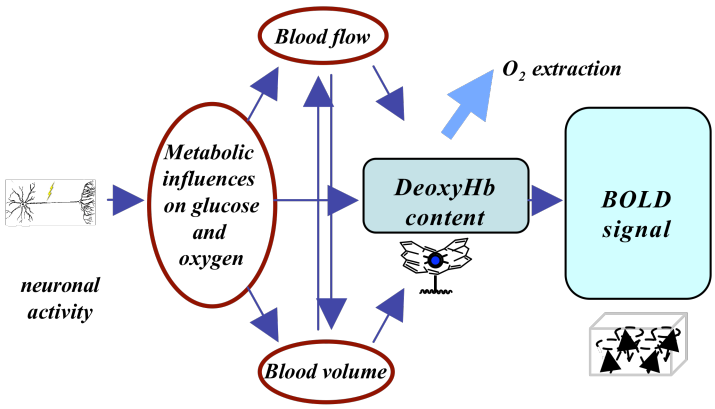

Fig. 3. Diagram showing the physiological influences on the BOLD signal.

inappropriate for single subject clinical studies (when false negatives could mislead the surgeon). Likewise, the standard "canonical" hemodynamic response function may be inappropriate in certain clinical populations where an altered temporal response may pertain. Indeed, there is good evidence that even in healthy subjects the hemodynamic response function changes with age (D'Esposito *et al*., 1999), most likely as the bio-mechanical properties of the vasculature changes. One way to minimise this potential problem is to use relatively long epochs of stimulation and control rather than event-related designs, such that the details of the hemodynamic response function do not dominate the shape of the model of neuronal response.

Physiological Implications in Clinical Application

There are additional potential problems that may be encountered when studying clinical populations. Clearly, as described above, the BOLD signal is a rather indirect method for assessment of neuronal activity. This is illustrated in Fig. 3, that shows the complex interaction between neuro-vascular signalling mechanisms, blood flow changes, blood volume changes, and tissue oxygen extraction effects that all contribute to the BOLD signal that is ultimately measured. This may mean that any disruptions to neurovascular coupling or hemodynamic response caused by the disease under study could result in misinterpretations in the BOLD response. Indeed, one could even anticipate the possibility of observing no BOLD effect at all, despite an underlying electrical activity, or even a paradoxical BOLD effect, that can only be interpreted in the light of other physiological information, such as corresponding measurements of the change in blood flow.

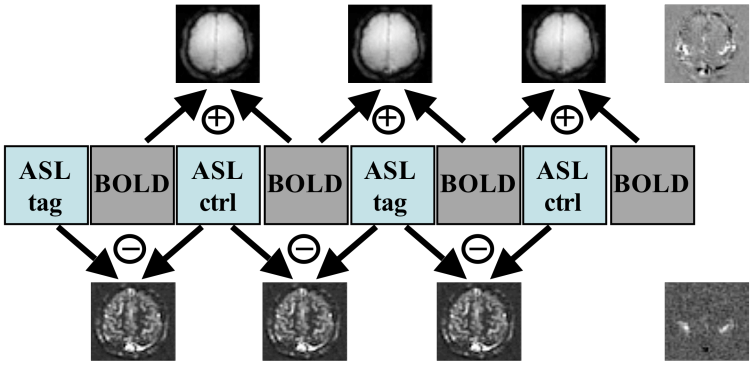

Fig. 4. Interlaced BOLD and ASL sequence allowing BOLD and perfusion data to be collected pseudo-simultaneously. Alternatively, a double echo ASL sequence can be used in which the second echo is BOLD weighted.

As implied above, there seems to be a good case for including in any clinical fMRI protocol a measurement of the changes in regional cerebral blood flow, and also ideally a measurement of resting blood flow (see Fig. 4). Indeed, there is increasing evidence that the BOLD effect is affected by shifts in the baseline level of cerebral blood flow, whereas the absolute change of regional cerebral blood flow is preserved (Brown *et al*., 2003). Also, there is good evidence that the BOLD effect can be either attenuated or augmented by the presence of vasoactive drugs (Seifritz *et al*. 2000; Mulderink *et al*. 2002) and even diet (Noseworthy *et al*. 2003), or can affect the temporal characteristics of the hemodynamic response (Liu *et al*. 2004). This clearly could have an impact in clinical research when comparing a group of medicated patients with a group of healthy control subjects.

Quality Assurance

An important aspect of performing fMRI in a clinical context is that there must be a well documented and rigorously applied procedure for assessment of scanner stability. Even then there may be differences in the fMRI measurements made at different sites, particularly when involving different field strengths (Zuo *et al*., 2005). It is therefore important to run regular quality assurance tests using a

phantom sample that loads the coil in a similar way to the human head (Stocker *et al.*, 2005; Friedman and Glover, 2006). Such data will reveal whether the system has stability problems that should be addressed before fMRI experiments are attempted. A recommended test that can be performed is a pseudo fMRI experiment in which a phantom sample (e.g. spherical agar gel phantom) is imaged using an otherwise standard fMRI protocol. One example would be an EPI time course consisting of 100 volumes, 64×64 pixels, 25 slices, TR/TE/thk=3000ms/40ms/5mm, FOV=24×24cm. A number of tests can then be performed on each slice of the data, including:

1) Mean signal variability: Variability in the time course of the mean signal from the central 80% of the image. This figure should reflect the transmitter stability and be better than 0.1%.

2) Signal-to-noise of the individual images: Signal=mean of central 20% of phantom, Noise=standard deviation of non-ghosted noise region. Ideally should be > 250. (Note that a true SNR measure requires a correction factor of 1.53 (Weisskoff, 1996) to account for the Rician nature of noise in magnitude MRI data).

3) Temporal signal-to-noise: For any two images, select ROI as central 20% of phantom. Signal=mean of ROI, Noise=standard deviation of ROI{image 1−image 2}. S/N result should be > 250.

4) Level of EPI ghost: For EPI sequences the ghost should be <2–3% of main image intensity.

5) Pixel-by-pixel temporal stability: For each pixel the temporal standard deviation in intensity is calculated. This is expressed as a percentage of the image mean intensity. Figure should be <0.5%.

Such tests need to be regularly run in order to be assured that clinical results are not compromised by unknown scanner instabilities.

Summary

Functional magnetic resonance imaging is a powerful technique with the ability to study brain function in the normal and diseased brain. However, there are a number of potential confounds when applying fMRI to clinical populations. For fMRI to have a role in clinical practice it is essential that standard protocols be developed and adopted, standard acquisition and data analysis be used, and regular quality assurance be practised. Even in clinical research it is important to be aware of potential confounds that may pertain when activation-flow coupling is affected by the disease under study. Finally, it is likely that the full clinical potential of BOLD fMRI will only be realised by combining it with other methods, such as perfusion MRI, diffusion tractography, transcranial magnetic stimulation, and electroencephalography.

References

Boxerman JL, Bandettini PA, Kwong KK, Baker JR, Davis TL, Rosen BR, Weisskoff RM. The intravascular contribution to fMRI signal change: Monte Carlo modeling and diffusion-weighted studies in vivo. Magn Reson Med. 1995 34(1):4-10.

Brown GG, Eyler Zorrilla LT, Georgy B, Kindermann SS, Wong EC, Buxton RB. BOLD and perfusion response to finger-thumb apposition after acetazolamide administration: differential relationship to global perfusion. J Cereb Blood Flow Metab. 2003 23(7):829-37.

Cohen ER, Ugurbil K, Kim SG. Effect of basal conditions on the magnitude and dynamics of the blood oxygenation level-dependent fMRI response. J Cereb Blood Flow Metab. 2002 22(9):1042-53.

D'Esposito M, Zarahn E, Aguirre GK, Rypma B. The effect of normal aging on the coupling of neural

activity to the bold hemodynamic response. Neuroimage. 1999 10(1):6-14.

de Graaf RA, Mason GF, Patel AB, Rothman DL, Behar KL. Regional glucose metabolism and glutamatergic neurotransmission in rat brain in vivo. Proc Natl Acad Sci U S A. 2004 101(34):12700-5. Epub 2004

Friedman, J, Glover GH, Quality assurance in fMRI data. J Magn Reson Imag. 2006 in press.

Friston KJ, Williams S, Howard R, Frackowiak RS, Turner R. Movement-related effects in fMRI time-series. Magn Reson Med. 1996 35(3):346-55.

Hutton C, Bork A, Josephs O, Deichmann R, Ashburner J, Turner R. Image distortion correction in fMRI: A quantitative evaluation. Neuroimage. 2002 16(1):217-40.

Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage. 2002 17(2):825-41.

Jezzard P, Balaban RS. Correction for geometric distortion in echo planar images from B0 field variations. Magn Reson Med. 1995 34(1):65-73.

Liu TT, Behzadi Y, Restom K, Uludag K, Lu K, Buracas GT, Dubowitz DJ, Buxton RB. Caffeine alters the temporal dynamics of the visual BOLD response. Neuroimage. 2004 23(4):1402-13.

Lu H, Golay X, Pekar JJ, Van Zijl PC. Functional magnetic resonance imaging based on changes in vascular space occupancy. Magn Reson Med. 2003 50(2):263-74.

Mulderink TA, Gitelman DR, Mesulam MM, Parrish TB. On the use of caffeine as a contrast booster for BOLD fMRI studies. Neuroimage. 2002 15(1):37-44.

Noseworthy MD, Alfonsi J, Bells S. Attenuation of brain BOLD response following lipid ingestion. Hum Brain Mapp. 2003 20(2):116-21.

Seifritz E, Bilecen D, Hanggi D, Haselhorst R, Radu EW, Wetzel S, Seelig J, Scheffler K. Effect of ethanol on BOLD response to acoustic stimulation: implications for neuropharmacological fMRI. Psychiatry Res. 2000 10;99(1):1-13.

Stocker T, Schneider F, Klein M, Habel U, Kellermann T, Zilles K, Shah NJ. Automated quality assurance routines for fMRI data applied to a multicenter study. Hum Brain Mapp. 2005 Jun;25(2):237-46.

Weisskoff RM. Simple measurement of scanner stability for functional NMR imaging of activation in the brain. Magn Reson Med. 1996 36(4):643-5.

Wong EC, Luh WM, Liu TT. Turbo ASL: arterial spin labeling with higher SNR and temporal resolution. Magn Reson Med. 2000 44(4):511-5.

Woods RP, Grafton ST, Holmes CJ, Cherry SR, Mazziotta JC. Automated image registration: I. General methods and intrasubject, intramodality validation. J Comput Assist Tomogr. 1998 22(1):139-52.

Zou KH, Greve DN, Wang M, Pieper SD, Warfield SK, White NS, Manandhar S, Brown GG, Vangel MG, Kikinis R, Wells WM 3rd. Reproducibility of Functional MR Imaging: Preliminary Results of Prospective Multi-institutional Study Performed by Biomedical Informatics Research Network. Radiology. 2005 237(3):781-9.